



별첨 사본은 아래 출원의 원본과 동일함을 증명함.

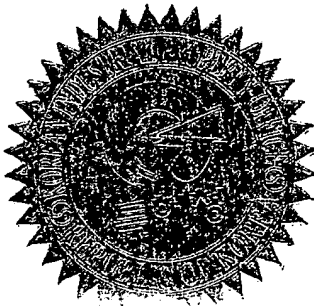
This is to certify that the following application annexed hereto  
is a true copy from the records of the Korean Industrial  
Property Office.

Best Available Copy

출원 번호 : 특허출원 2000년 제 86284 호  
Application Number

출원 년 월 일 : 2000년 12월 29일  
Date of Application

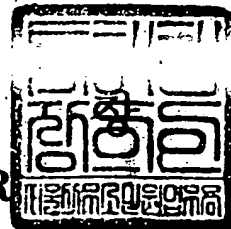
출원인 : 삼성전자 주식회사 외 1명  
Applicant(s)



2001 년 02 월 19 일

특 허 청

COMMISSIONER



【서류명】	특허출원서
【권리구분】	특허
【수신처】	특허청장
【참조번호】	0025
【제출일자】	2000.12.29
【국제특허분류】	G06F
【발명의 명칭】	객체 활동 모델링 방법
【발명의 영문명칭】	Object activity modeling method
【출원인】	
【명칭】	삼성전자 주식회사
【출원인코드】	1-1998-104271-3
【출원인】	
【명칭】	더 리전트 오브 더 유니버시티 오브 캘리포니아
【출원인코드】	5-1999-020685-1
【대리인】	
【성명】	이영필
【대리인코드】	9-1998-000334-6
【포괄위임등록번호】	1999-009556-9
【포괄위임등록번호】	2000-055014-6
【대리인】	
【성명】	최흥수
【대리인코드】	9-1998-000657-4
【포괄위임등록번호】	1999-009578-0
【포괄위임등록번호】	2000-056092-0
【대리인】	
【성명】	이해영
【대리인코드】	9-1999-000227-4
【포괄위임등록번호】	2000-002816-9
【포괄위임등록번호】	2000-057184-1
【발명자】	
【성명의 국문표기】	최양림
【성명의 영문표기】	CHOI, Yang Lim
【주민등록번호】	710120-1830615
【우편번호】	442-192

**【주소】** 경기도 수원시 팔달구 우만2동 105 우만선경아파트 102동 1112호  
**【국적】** KR  
**【발명자】**  
**【성명의 국문표기】** 유윤주  
**【성명의 영문표기】** YU, Yun Ju  
**【주민등록번호】** 710725-1836311  
**【우편번호】** 609-391  
**【주소】** 부산광역시 금정구 장전1동 130-23 18/3  
**【국적】** KR  
**【발명자】**  
**【성명의 국문표기】** 비 .에스.만주나스  
**【성명의 영문표기】** B.S.,Manjunath  
**【주소】** 미합중국 캘리포니아 93106-9560 산타바바라, 유니버시티 오브 캘리 포티아  
**【국적】** US  
**【발명자】**  
**【성명의 국문표기】** 첸 칭-에이  
**【성명의 영문표기】** CHEN, Ching-Wei  
**【주소】** 미국 캘리포니가 93117 아이슬라비스타 6503 마드리드 알 디 아파트 제4호  
**【국적】** CN  
**【우선권주장】**  
**【출원국명】** US  
**【출원종류】** 특허  
**【출원번호】** 60/248,010  
**【출원일자】** 2000.11.14  
**【증명서류】** 미첨부  
**【심사청구】** 청구  
**【취지】** 특허법 제42조의 규정에 의한 출원, 특허법 제60조의 규정에 의한 출원심사를 청구합니다. 대리인  
 이영필 (인) 대리인  
 최흥수 (인) 대리인  
 이해영 (인)

**【수수료】**

【기본출원료】 20 면 29,000 원

【가산출원료】 1 면 1,000 원

【우선권주장료】 1 건 26,000 원

【심사청구료】 11 항 461,000 원

【합계】 517,000 원

**【첨부서류】**

1. 요약서·명세서(도면)\_1통 2.우선권증명서류 및 동 번역문\_1통[추후제출]

**【요약서】****【요약】**

인간 활동과 같이 복잡한 객체의 활동을 효율적으로 분석하는 객체 활동 모델링 방법이 개시된다. 본 발명에 따른 객체 활동 모델링 방법은 (a) 영상 시퀀스로부터 옵티컬 플로우 벡터를 구하는 단계와, (b) 옵티컬 플로우 벡터를 사용하여 복수 개의 영상 프레임(video frame)들에 대한 특징 벡터의 확률 분포를 구하는 단계와, (c) 특징 벡터의 확률 분포를 사용하여 상태들을 모델링하는 단계, 및 (d) 상태 전환을 기초로 영상 시퀀스내의 객체의 활동을 묘사하는 단계를 포함하는 것을 특징으로 한다. 본 발명에 따르면, 비디오 인덱싱 및 인식 분야에서 인간 활동과 같은 복잡한 활동을 객체의 분할 없이 효과적으로 모델링하고 인식하는 것이 가능하다.

**【대표도】**

도 1

**【명세서】****【발명의 명칭】**

객체 활동 모델링 방법{Object activity modeling method}

**【도면의 간단한 설명】**

도 1은 본 발명의 실시예에 따른 객체 활동 모델링 방법의 주요 단계들을 나타낸 흐름도이다.

도 2a는 사람이 일어서기 시작하다가 다시 앉는 활동에 대한 학습전의 히든마코브 모델의 예를 나타낸 도면이다.

도 2b는 사람이 일어서기 시작하다가 다시 앉는 활동에 대한 학습후의 히든마코브 모델의 예를 나타낸 도면이다.

도 3은 본 발명의 실시예에 따른 객체 활동 인식 방법의 주요 단계들을 나타낸 흐름도이다.

**【발명의 상세한 설명】****【발명의 목적】****【발명이 속하는 기술분야 및 그 분야의 종래기술】**

<5> 본 발명은 객체 활동 모델링 방법에 관한 것으로, 더 상세하게는 인간 활동과 같이 복잡한 객체의 활동을 효율적으로 분석하는 객체 활동 모델링 방법에 관한 것이다. 또한, 본 발명은 객체 활동 모델링 방법에 의하여 모델링된 객체 활동 모델을 사용하여 비디오 시퀀스내의 객체의 활동 또는 이벤트를 인식할 수 있는 객체 활동 인식 방법에 관한 것이다.

<6> 인간이 앉거나, 서거나, 걷거나, 또는 몸을 돌리는 등과 같은 인간의 활동은 카메라를 사용하여 캡처되어 디지털 비디오로써 저장될 수 있다. 디지털 비디오를 얻은 다음에는 그 내용을 분석하는 것이 가능하다. 예를들어, 상기 디지털 비디오의 활동의 시간적 또는 공간적 특징들은 트레이닝 데이터를 기반으로 하는 통계적인 모델들을 사용하여 묘사(characterize)될 수 있다. 이러한 모델들은 패턴 인식을 위하여 주어진 비디오 시퀀스를 데이터베이스 영상들에 정합시키는데 사용될 수 있다. 패턴 분석 다음에는 비디오를 이러한 패턴들을 사용하여 인지적으로(semantically) 인덱싱할 수 있다. 또한, 이 과정에서 비디오 콘텐츠의 인지적인 요약물을 얻을 수 있다.

<7> 종래의 객체 활동 분석 방법은 두가지 부류로 분류될 수 있다. 첫 번째 부류의 방법에서는, 활동을 분석하도록 설계된 장치가 인간에게 부착되어 그 장치가 활동 분석에 사용된다. 두 번째 부류의 방법에서는, 객체의 그림 또는 기하학적 특징들이 활동 분석에 사용된다. 하지만, 첫 번째 부류의 방법에서는, 인간에게 장치를 부착하여야 하기 때문에, 인간에게 부착된 장치들이 인간의 활동을 제한한다는 문제점이 있다. 또한, 두 번째 부류의 방법은 영상으로부터 개별 객체들을 분할(segmentation)할 것이 요구된다. 하지만, 영상으로부터 개별 객체들을 정확히 분할할 수 없는 경우가 많다. 특히, 분할이 어려운 인체와 같은 복잡한 객체들에서는 적용하기가 어렵다는 문제점이 있다.

#### 【발명이 이루고자 하는 기술적 과제】

<8> 본 발명이 이루고자 하는 기술적 과제는 인체와 같이 복잡한 객체들을 모델링할 수 있는 객체 활동 모델링 방법을 제공하는 것이다.

<9> 본 발명이 이루고자 하는 또 다른 기술적 과제는 상기와 같은 객체 활동 모델링 방

법에 의하여 모델링된 활동 모델을 사용하여 객체의 활동을 인식하는 방법을 제공하는 것이다.

【발명의 구성 및 작용】

- <10> 상기 과제를 이루기 위하여 본 발명에 따른 객체 활동 모델링 방법은 (a) 영상 시퀀스로부터 옵티컬 플로우 벡터를 구하는 단계; (b) 옵티컬 플로우 벡터를 사용하여 복수 개의 영상 프레임들에 대한 특징 벡터의 확률 분포를 구하는 단계; (c) 특징 벡터의 확률 분포를 사용하여 상태들을 모델링하는 단계; 및 (d) 상태 전환을 기초로 영상 시퀀스내의 객체의 활동을 묘사하는 단계;를 포함하는 것을 특징으로 한다.
- <11> 또한, 상기 (a) 단계는 어파인 움직임 추정을 기초로 하는 것이 바람직하다.
- <12> 또한, 상기 (a) 단계는, (a-1) 입력된 영상 프레임들을 복수 개의 영상 프레임 그룹들로 그룹핑하고 그룹핑된 각 영상 프레임 그룹들을 개별 상태로써 구분하는 단계; (a-2) 각 개별 상태의 영상 프레임 그룹내의 각 영상별로 움직임 추정에 의하여 어파인 움직임 변수를 구하는 단계; 및 (a-3) 어파인변수들로부터 옵티컬 플로우 벡터를 구하는 단계;를 포함하는 것이 바람직하다.
- <13> 또한, 상기 (a-2) 단계는,  $I$ 는 밝기(intensity),  $t$ 는 시간,  $x$ 는 픽셀 위치  $(x, y)$ ,  $v$ 는 움직임 벡터라 할 때,  $I_t(x) = I_{t-1}(x - V(x))$ 로써 나타내어지는 객체상의 픽셀의 밝기를 기반으로 하여, 주어진 영상에 걸친 제곱차합(summed square difference)  $\sum (I_t(x) - I_{t-1}(x - v(x)))^2$ 을 최소화하는 변수들을 움직임 변수로써 결정하는 단계;를 포함하는 것이 바람직하다.
- <14> 또한, 상기 (b) 단계는,  $P = (p_1, p_2, \dots, p_d)$ 는 각 픽셀 위치  $(x, y)$ 에서 계산된 움직임



변수,  $L$ 은 영상 프레임(video frame) 또는 관심 영역내의 픽셀의 수,  $d$ 는 차원의 수,  $d \times L$ 차원 벡터인 특징 벡터  $Z$ 는  $Z = (p_1^1, p_1^2, \dots, p_1^L, p_2^1, p_2^2, \dots, p_2^L, p_d^1, p_d^2, \dots, p_d^L)^T$ ,  $m$ 은 특징 벡터  $Z$ 의 평균 벡터,  $Q$ 는 특징 벡터  $Z$ 의 코베리언스 행렬(covariance matrix)이라 할 때, 관찰 클래스(observation class)  $\Omega$ 로부터 특징 벡터  $Z$ 가 주어지면, 확률 분포  $P(Z|\Omega)$ 를

$$P(Z|\Omega) = \frac{\exp(-\frac{1}{2}(Z-m)^T Q^{-1}(Z-m))}{(2\pi)^N |Q|^{1/2}}$$

에 따라 계산하는 단계;를 포함하는 것이 바람직하다.

<15> 또한, 상기 (b) 단계는,  $\hat{Z} = Z - m$ ,  $\Phi$ 의 열은 코베리언스 행렬  $Q$ 의 직교 고유 벡터(orthonormal eigenvectors),  $\Lambda$ 는  $Q$ 의 대각 고유값(diagonal eigenvalue)이라 할 때, 코베리언스 행렬  $Q$ 를  $Q = \Phi \Lambda \Phi^T$ 로써 분해하는 단계;  $M$ 은 주요 성분의 수,  $y_i$ 는 투영 계수  $Y$ 의  $i$  번째 성분,  $\alpha_i$ 는  $Q$ 의  $i$  번째 고유값,  $\rho$ 는  $\rho = \frac{1}{N-M} \sum_{i=1}^M \alpha_i$ 에 의하여 구해지는 최적값이라 할 때, 관찰 클래스(observation class)  $\Omega$ 로부터 특징 벡터  $Z$ 가 주어지면, 확률 분포  $P(Z|\Omega)$ 를

$$P(Z|\Omega) = \left[ \frac{\exp(-\frac{1}{2} \sum_{i=1}^M y_i^2 / \alpha_i)}{(2\pi)^M |\Lambda|^{1/2}} \right] \left[ \frac{\exp(-\frac{1}{2} \sum_{i=M+1}^N y_i^2 / 2\rho)}{(2\pi)^{(N-M)/2}} \right]$$

에 따라 계산하는 단계;를 포함하는 것이 바람직하다.

<16> 또한, 상기 (c) 단계는, 상태 전환을 기초로 영상 시퀀스내의 객체의 활동을 히든 마코브모델(Hidden Markov Model: HMM)을 사용하여 묘사하는 단계;를 포함하는 것이 바람직하다.

<17> 또한, 상기 히든마코브모델(HMM)은,  $N$ 은 가능한 상태의 수,  $\Xi = \{q_1, q_2, \dots, q_N\}$ ,  $A$ 는 은닉된 상태  $i$ 와  $j$  사이의 전환  $\{a_{ij}\}$ ,  $B$ 는 상태  $j$ 에 해당하는 관측 심볼 확률  $\{b_j(\cdot)\}$ ,  $\pi$ 는 초기 상태 분포라 할 때,  $\lambda = \{\Xi, A, B, \pi\}$ 로써 표현되고, 상태  $\Xi = \{q_1, q_2, \dots, q_N\}$ 와 초기 상태 분포  $\pi$ 는 비디오 데이터를 기초로 미리 결정되는 것이 바람직하다.

- <18> 또한, 상기 다른 과제를 이루기 위하여 본 발명에 따른 객체 활동 인식 방법은 (a) 영상 프레임들에 대한 움직임 추정에 의하여 특징 벡터들을 구하는 단계; (b) 구한 특징 벡터들을 사용하여 각 프레임이 속하는 상태를 결정하는 단계; 및 (c) 결정된 상태에 대하여 전환 행렬을 사용하여 주어진 활동 모델 사전으로부터 주어진 영상 프레임과 활동 모델사이의 유사도를 최대화하는 하나의 활동 모델을 인식된 활동으로써 결정하는 단계;를 포함하는 것을 특징으로 한다.
- <19> 이하 첨부된 도면들을 참조하여 본 발명의 바람직한 실시예들을 상세히 설명하기로 한다.
- <20> 도 1에는 본 발명의 실시예에 따른 객체 활동 모델링 방법의 주요 단계들을 흐름도로써 나타내었다. 모든 종류의 객체 활동들은 각기 다른 스타일의 움직임 변화(motion changes)중에 있는 객체들로 해석될 수 있으므로, 활동들을 객체의 움직임 분포와 연계시켜 고려하는 것이 바람직하다. 따라서, 본 발명에서는 움직임 분포를 기초로 인간 활동을 모델링한다. 본 실시예에서는 정밀한 움직임 추정을 적용하지 않고, 모델 기반의 움직임 추정을 적용한다.
- <21> 도 1을 참조하면, 본 발명의 실시예에 따른 객체 활동 모델링 방법에서는 먼저, 수작업에 의한 상태 모델 선택(Manual state model selection)을 사용하여 입력된 영상 프레임들을 복수 개의 영상 프레임 그룹들로 그룹핑하고 그룹핑된 각 영상 프레임 그룹들을 개별 상태로써 구분한다(단계 102).
- <22> 각 개별 상태의 영상 프레임 그룹내의 각 영상별로 어파인 움직임 추정에 의하여 어파인 움직임 변수를 구한다(단계 104). 여기서, 움직임 추정은,  $I$ 는 밝기(intensity),  $t$ 는 시간,  $x$ 는 픽셀 위치 ( $x, y$ ),  $v$ 는 움직임 벡터라 할 때,

## &lt;23&gt; 【수학식 1】

$$I_t(x) = I_{t-1}(x - V(x))$$

<24> 로써 나타내어지는 객체상의 픽셀의 밝기를 기반으로 한다. 즉, 주어진 영역에 걸친 다음의 제곱차합(summed square difference)  $\sum (I_t(x) - I_{t-1}(x - v(x)))^2$ 을 최소화하는 변수들이 움직임 변수로써 추정된다.

<25> 모델 기반의 움직임 추정에서는, 객체의 크기가 카메라와 객체 사이의 거리에 비하여 훨씬 적으면, 객체의 움직임은 어파인 모델(affine model)을 사용하여 근사화될 수 있다. 영상내의 각 점을 예를들어  $5 \times 5$  픽셀크기와 같이 작은 국부적 윈도우를 사용하여 나타내면, 어파인 모델 변수를 사용하여 움직임을 근사화할 수 있다. 어파인 모델은,

## &lt;26&gt; 【수학식 2】

$$V(x,y) = \Psi(x,y)K$$

<27> 과 같이 나타내어 진다. 여기서,  $(x,y)$ 는 객체상의 임의의 한 점의 좌표,

$V(x,y) = (u(x,y), v(x,y))^T$ 는 움직임 벡터,  $K = (k_1, k_2, k_3, k_4, k_5, k_6)^T$ 는 어파인 모델 변수,

$\Psi(x,y) = \begin{pmatrix} 1 & x & y & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & x & y \end{pmatrix}$ 이다. 또한,  $k_1$ 과  $k_4$ 는 직교 이동에 해당하고,  $k_2$ ,  $k_3$ ,  $k_5$ , 및

$k_6$ 는 표면의 변형(deformation)에 해당한다는 것에 주목하여야 한다. 이때,  $k_2$ ,  $k_3$ ,  $k_5$ , 및  $k_6$ 를 무시하면, 움직임 벡터  $V$ 는  $V = (k_1, k_4)^T$ 로써 나타낼 수 있다. 이러한 움

직임 벡터  $V = (k_1, k_4)^T$ 는 전형적인 옵티컬 플로우 벡터(optical flow vector)이다. 따라서, 어파인 움직임 변수들로부터 옵티컬 플로우 벡터를 구할 수 있다(단계 106).

<28> 이제, 각 픽셀 위치  $(x,y)$ 에서 계산된 움직임 변수  $P = (p_1, p_2, \dots, p_d)$ 를 고려한다. 예를들어,  $P$ 는 6차원 어파인 움직임 변수 또는 2차원 옵티컬 플로우벡터일 수 있다. 이러

한, 변수값들은,  $L$ 은 영상 프레임 또는 관심 영역내의 픽셀의 수,  $d$ 는 차원의 수라 할 때,

<29> 【수학식 3】

$$Z = (p_1^1, p_1^2, \dots, p_1^L, p_2^1, p_2^2, \dots, p_2^L, p_d^1, p_d^2, \dots, p_d^L)^T$$

<30> 즉, 어파인 움직임 변수들 또는 옵티컬 플로우 벡터로 이루어지는 특징 벡터  $Z$ 는  $d \times L$ 차원 벡터이다. 따라서, 옵티컬 플로우 벡터들을 특징 벡터  $Z$ 를 얻는다(단계 108).

<31> 한편, 특징 벡터  $Z$ 를 가우시안 함수로써 모델링할 수 있다. 이때, 상기 가우시안 함수의 평균을  $m$ , 행렬로 나타내어지는 특징 벡터  $Z$ 의 코베리언스 행렬을  $Q$ 라 나타낸다. 이제, 관찰 클래스(observation class)  $\Omega$ 로부터 특징 벡터  $Z$ 가 주어지면, 확률 분포  $P(Z | \Omega)$ 는,

<32> 【수학식 4】

$$P(Z | \Omega) = \frac{\exp(-\frac{1}{2} (Z-m)^T Q^{-1} (Z-m))}{(2\pi)^N |Q|^{1/2}}$$

<33> 과 같이 계산될 수 있다. 여기서,  $Z$ 는 특징 벡터,  $m$ 은 특징 벡터  $Z$ 의 평균 벡터,  $Q$ 는 특징 벡터  $Z$ 의 코베리언스 행렬(covariance matrix)를 나타낸다.

<34> 하지만, 상기 수학식 4에 따라 관측 클래스에 대하여 확률을 계산하면, 영상의 픽셀수와 차원의 수를 고려하면, 확률의 계산량이 매우 많다. 따라서, 본 실시예에서는 카르넨-뢰베 변환(KLT: Karhunen-Loeve Transform)을 사용하여 상기 수학식 4의 계산을 단순화한다. 먼저,  $\hat{Z} = Z - m$ 라 둔다. 다음으로, 코베리언스 행렬은,  $\Phi$ 의 열은  $Q$ 의 직교 고유 벡터(orthonormal eigenvectors),  $\Lambda$ 는  $Q$ 의 대각 고유값(diagonal eigenvalue)이라 할 때,

## &lt;35&gt; 【수학식 5】

$$Q = \Phi \Lambda \Phi^T$$

<36> 로써 분해될 수 있다. 이를 기초로,  $M$ 은 주요 성분의 수,  $y_i$ 는 투영 계수  $Y$ 의  $i$  번째 성분,  $\alpha_i$ 는  $Q$ 의  $i$  번째 고유값,  $\rho$ 는  $\rho = \frac{1}{N-M} \sum_{i=1}^M \alpha_i$ 에 의하여 구해지는 최적값이라 할 때, 상기 수학식 4는,

## &lt;37&gt; 【수학식 6】

$$P(Z | \Omega) = \left[ \frac{\exp(-\frac{1}{2} \sum_{i=1}^M y_i^2 / \alpha_i)}{(2\pi)^M |\Lambda|^{1/2}} \right] \left[ \frac{\exp(-\frac{1}{2} \sum_{i=M+1}^N y_i^2 / 2\rho)}{(2\pi\rho)^{(N-M)/2}} \right]$$

<38> 으로서 근사화될 수 있다. 따라서, 본 실시예에서는 관찰 클래스(observation class)  $\Omega$ 로부터 특징 벡터  $Z$ 가 주어지면, 확률 분포  $P(Z | \Omega)$ 를 상기 수학식 6을 사용하여 계산한다(단계 110). 다음으로, 상기와 같이 계산된 확률 분포를 사용하여 각 상태들을 모델링한다(단계 112).

<39> 다음으로, 상태 전환을 기초로 영상 시퀀스내의 객체의 활동을 묘사한다(단계 114). 본 실시예에서는 영상 시퀀스 내의 객체의 활동을 묘사하기 위하여 히든마코브모델(Hidden Markov Model: HMM)을 사용한다. 히든마코브모델은 시간에 따라 다양하게 변화하는 자료들을 학습하고 인식할 때 성능이 우수한 통계적 모델로 널리 알려져 있다. 특히, 히든 마코브 모델은 연속적으로 입력되는 온라인 문자 또는 음성의 인식에 널리 사용된다. 히든 마코브모델을 사용한 음성 인식에서는, 음성이 마코브 모델로 모델링될 수 있다는 가정하에, 트레이닝 과정에서 마코브 모델의 확률적인 변수를 구하여 기준 마코브 모델을 생성한다. 또한, 음성 인식 과정에서는 입력음성과 가장 유사한 기준 마코브 모델을 추정함으로써 음성을 인식한다. 일반적으로, 음성인식을 위한 모델로써 히든

마코브 모델(hidden Markov model)을 사용하는데 이것은 음성패턴의 다양한 변화를 수용하기 위해서이다. 여기서 히든(hidden)이라는 용어는 상태가 음성 패턴에 관계없이 모델 속에 숨어있다는 것을 뜻한다. 일반적인 히든마코브모델(HMM)은,  $N$ 은 가능한 상태의 수,  $\Xi$ 는  $\Xi = \{q_1, q_2, \dots, q_N\}$ ,  $A$ 는 은닉된 상태  $i$ 와  $j$  사이의 전환  $\{a_{ij}\}$ ,  $B$ 는 상태  $j$ 에 해당하는 관측 심볼 확률  $\{b_j(\cdot)\}$ ,  $\pi$ 는 초기 상태 분포라 할 때,

<40> 【수학식 7】

$$\lambda = \{\Xi, A, B, \pi\}$$

<41> 로써 표현될 수 있다. 상태  $\Xi = \{q_1, q_2, \dots, q_N\}$ 와 초기 상태 분포  $\pi$ 는 비디오 데이터를 기초로 미리 결정된다. 히든 마코브 모델 변수  $A$ 와  $B$ 는 알려진 바움-웰쉬(Baum-Welsh) 재추정 공식을 사용하여 반복적으로 트레이닝될 수 있다.

<42> 상태 또는 상태 모델의 수는 경험적으로 결정될 수 있으나, 본 실시예에서는 4개의 상태를 선택한 것을 예로써 설명한다. 또한, 본 실시예에서는 4개의 상태를 사용하여 활동을 모델링하며, 각 상태의 전환 확률에 동일한 값을 주는 것을 예로써 설명한다.

<43> 도 2a 및 도 2b에는 일어서기 시작하다가 다시 앉는 활동(이하 bd라 칭함)에

대한 히든마코브모델의 예를 나타내었다. 도 2a에는 학습전의 bd에 대한 히든마코브모델의 예를 나타내었고, 도 2b에는 학습후의 bd에 대한 히든마코브모델의 예를 나타내었다. 도 2a를 참조하면, 하나의 상태에서 다른 하나의 상태로 전환하거나 이전의 상태로 복귀할 확률이 동일하게 0.333으로 동일하게 설정되어 있다. 4로 참조되는 상태에서 4로 복귀되는 확률은 모델 전개의 편의상 1로 놓는다. 하지만, 도 2b를 참조하면, 하나의 상태에서 다른 하나의 상태로 전환하거나 이전의 상태로 복귀할 확률이 각기 다르게 설정되어 있다. 다음으로, 다르게 설정된 전환 확률을 사용하여 전환 행렬을 구한다. 이제, 각기 다른 확률 분포로써 정의되는 서로 다른 복수 개의 상태들과 구해진 전환 행렬을 활동 모델로써 결정한다. 이로써, 활동에 대한 모델링이 완료된다.

<44>       상기와 같은 객체 활동 모델링 방법에 따르면, 비디오 인덱싱 및 인식 분야에서 인간 활동과 같은 복잡한 활동을 효율적으로 모델링할 수 있다. 특히, 상기 객체 활동 모델링 방법에 따르면, 활동 인식을 위하여 필요한 객체 활동을 객체의 분할없이 모델링하는 것이 가능하다.

<45>       이상의 실시예에 의한 객체 활동 모델링 방법은 정적 카메라(static cameras)와 같은 시스템에 적용할 수 있다. 하지만, 적용하고자 하는 시스템이 모션 카메라(camera motion)인 경우에는 인간 움직임을 먼저 복구하여야 한다. 이후의 과정들은 이상의 실시예에서와 동일하다.

<46>       이제, 활동을 인식하는 과정들을 기술한다. 도 3에는 본 발명의 실시예에 따른 객체 활동 인식 방법의 주요 단계들을 흐름도로써 나타내었다. 도 3을 참조하면, 먼저, 활동을 인식하고자 하는 영상 프레임들을 입력한다(단계 302). 다음으로, 입력된 영상 프레임들에 대한 움직임 추정에 의하여 특징 벡터들을 구한다(단계 304). 단계(304)는 도

1을 참조하여 설명한 단계(106)과 실질적으로 동일하다고 이해될 수 있다.

<47> 다음으로, 구한 특징 벡터들을 사용하여 각 프레임이 속하는 상태를 결정(단계 306)하고, 결정된 상태에 대하여 전환 행렬을 사용하여, T는 비디오 시퀀스를 구성하는 프레임의 수를 나타내는 양의 정수,  $Z_1, Z_1, \dots, Z_T$ 는 각각 1, 2, ..., T 번째 프레임의 특징 벡터들이라 할 때, 영상 프레임  $O = \{Z_1, Z_1, \dots, Z_T\}$ 이 주어지면, E는 상태 모델의 수라 할 때, 주어진 활동 모델 사전  $\{\lambda_1, \lambda_2, \dots, \lambda_E\}$ 으로부터 주어진 영상 프레임과 활동 모델사이의 유사도  $P(O|\lambda)$ 를 최대화하는 하나의 활동 모델을 인식된 활동으로써 결정한다(단계 308). 전환 행렬은 트레이닝 과정에서 장면 j에 해당하는 관측 심볼 확률  $B = \{b_j(i)\}$ 을 기초로 EM(expectation-maximization) 알고리즘을 사용하여 얻는다. 탐색 속도를 향상시키기 위해서는 초기화, 상태 예측, 및 측정 갱신의 세 단계로 이루어지는 칼만 필터(Kalman filter)의 예측 알고리즘을 기초로 트레이닝시에 사용된 크기와 동일한 크기를 가지는 윈도우 내의 이동 궤적을 추적하는 것이 보다 바람직하다.

<48> 이로써, 영상 시퀀스내의 인간과 같은 복잡한 객체의 활동을 인식하는 것이 가능하다. 상기와 같은 객체 활동 인식 방법에 따르면, 인간의 활동과 같이 복잡한 활동을 효과적으로 인식한다. 특히, 상기 객체 활동 인식 방법에 따르면, 객체를 분할하지 않고도 활동을 인식할 수 있다.

<49> 또한, 상기와 같은 본 발명에 따른 객체 활동 모델링 방법과 객체 활동 인식 방법은 개인용 또는 서버급의 컴퓨터내에서 실행되는 프로그램으로 작성 가능하다. 상기 프로그램을 구성하는 프로그램 코드들 및 코드 세그먼트들은 당해 분야의 컴퓨터 프로그래머들에 의하여 용이하게 추론될 수 있다. 또한, 상기 프로그램은 컴퓨터 독취 가능 기록



매체에 저장될 수 있다. 상기 기록 매체는 자기기록매체, 광기록 매체, 및 전파 매체를 포함한다.

【발명의 효과】

<50> 상술한 바와 같이 본 발명에 따르면 비디오 인덱싱 및 인식 분야에서 인간 활동과 같은 복잡한 활동을 객체의 분할없이 모델링하고 인식하는 것이 가능하다.

**【특허청구범위】****【청구항 1】**

- (a) 영상 시퀀스로부터 옵티컬 플로우 벡터를 구하는 단계;
- (b) 옵티컬 플로우 벡터를 사용하여 복수 개의 영상 프레임들에 대한 특징 벡터의 확률 분포를 구하는 단계;
- (c) 특징 벡터의 확률 분포를 사용하여 상태들을 모델링하는 단계; 및
- (d) 상태 전환을 기초로 영상 시퀀스내의 객체의 활동을 묘사하는 단계;를 포함하는 것을 특징으로 하는 객체 활동 모델링 방법.

**【청구항 2】**

제1항에 있어서, 상기 (a) 단계는,

어파인 움직임 추정을 기초로 하는 것을 특징으로 하는 객체 활동 모델링 방법.

**【청구항 3】**

제2항에 있어서, 상기 (a) 단계는,

- (a-1) 입력된 영상 프레임들을 복수 개의 영상 프레임 그룹들로 그룹핑하고 그룹핑된 각 영상 프레임 그룹들을 개별 상태로써 구분하는 단계;
- (a-2) 각 개별 상태의 영상 프레임 그룹내의 각 영상별로 움직임 추정에 의하여 어파인 움직임 변수를 구하는 단계; 및
- (a-3) 어파인변수들로부터 옵티컬 플로우 벡터를 구하는 단계;를 포함하는 것을 특징으로 하는 객체 활동 모델링 방법.

## 【청구항 4】

제3항에 있어서, 상기 (a-2) 단계는,

$I$ 는 밝기(intensity),  $t$ 는 시간,  $x$ 는 픽셀 위치 ( $x, y$ ),  $v$ 는 움직임 벡터라 할 때,  $I_t(x) = I_{t-1}(x-v(x))$ 로써 나타내어지는 객체상의 픽셀의 밝기를 기반으로 하여, 주어진 영상에 걸친 제곱차합(summed square difference)  $\sum (I_t(x) - I_{t-1}(x-v(x)))^2$ 을 최소화하는 변수들을 움직임 변수로써 결정하는 단계;를 포함하는 것을 특징으로 하는 객체 활동 모델링 방법.

## 【청구항 5】

제1항에 있어서, 상기 (b) 단계는,

$P = (p_1, p_2, \dots, p_d)$ 는 각 픽셀 위치 ( $x, y$ )에서 계산된 움직임 변수,  $L$ 은 영상 프레임 또는 관심 영역내의 픽셀의 수,  $d$ 는 차원의 수,  $d \times L$ 차원 벡터인 특징 벡터  $Z$ 는  $Z = (p_1^1, p_1^2, \dots, p_1^L, p_2^1, p_2^2, \dots, p_2^L, p_d^1, p_d^2, \dots, p_d^L)^T$ ,  $m$ 은 특징 벡터  $Z$ 의 평균 벡터,  $Q$ 는 특징 벡터  $Z$ 의 코베리언스 행렬(covariance matrix)이라 할 때, 관찰 클래스(observation class)  $\Omega$ 로부터 특징 벡터  $Z$ 가 주어지면, 확률 분포  $P(Z | \Omega)$ 를 
$$P(Z | \Omega) = \frac{\exp(-\frac{1}{2}(Z-m)^T Q^{-1}(Z-m))}{(2\pi)^N |Q|^{1/2}}$$
에 따라 계산하는 단계;를 포함하는 것을 특징으로 하는 객체 활동 모델링 방법.

## 【청구항 6】

제1항에 있어서, 상기 (b) 단계는,

$\hat{Z} = Z - m$ ,  $\Phi$ 의 열은 코베리언스 행렬  $Q$ 의 직교 고유 벡터(orthonormal eigenvectors),  $\Lambda$ 는  $Q$ 의 대각 고유값(diagonal eigenvalue)이라 할 때, 코베리언스 행렬  $Q$ 를  $Q = \Phi \Lambda \Phi^T$ 로써 분해하는 단계;

$M$ 은 주요 성분의 수,  $y_i$ 는 투영 계수  $Y$ 의  $i$  번째 성분,  $\alpha_i$ 는  $\Omega$ 의  $i$  번째 고유값,  $\rho = \frac{1}{N-M} \sum_{i=1}^N \alpha_i$ 에 의하여 구해지는 최적값이라 할 때, 관찰 클래스

(observation class)  $\Omega$ 로부터 특징 벡터  $Z$ 가 주어지면, 확률 분포  $P(Z|\Omega)$ 를

$$P(Z|\Omega) = \left[ \frac{\exp(-\frac{1}{2} \sum_{i=1}^M y_i^2 / \alpha_i)}{(2\pi)^M |\Lambda|^{1/2}} \right] \left[ \frac{\exp(-\frac{1}{2} \sum_{i=M+1}^N y_i^2 / 2\rho)}{(2\pi\rho)^{(N-M)/2}} \right]$$

에 따라 계산하는 단계를 포함하는 것을 특징으로 하는 객체 활동 모델링 방법.

#### 【청구항 7】

제1항에 있어서, 상기 (c) 단계는,

상태 전환을 기초로 영상 시퀀스내의 객체의 활동을 히든마코브모델(Hidden Markov Model: HMM)을 사용하여 묘사하는 단계;를 포함하는 것을 특징으로 하는 객체 활동 모델링 방법.

#### 【청구항 8】

제7항에 있어서,

상기 히든마코브모델(HMM)은,

$N$ 은 가능한 상태의 수,  $\Xi$ 는  $\Xi = \{q_1, q_2, \dots, q_N\}$ ,  $A$ 는 은닉된 상태  $i$ 와  $j$  사이의 전환  $\{a_{ij}\}$ ,  $B$ 는 상태  $j$ 에 해당하는 관측 심볼 확률  $\{b(\cdot)\}$ ,  $\pi$ 는 초기 상태 분포라 할 때,  $\lambda = \{\Xi, A, B, \pi\}$ 로써 표현되고,

상태  $\Xi = \{q_1, q_2, \dots, q_N\}$ 와 초기 상태 분포  $\pi$ 는 비디오 데이터를 기초로 미리 결정되는 것을 특징으로 하는 객체 활동 모델링 방법.

## 【청구항 9】

(a) 영상 프레임들에 대한 움직임 추정에 의하여 특징 벡터들을 구하는 단계;  
 (b) 구한 특징 벡터들을 사용하여 각 프레임이 속하는 상태를 결정하는 단계; 및  
 (c) 결정된 상태에 대하여 전환 행렬을 사용하여 주어진 활동 모델 사전으로부터 주어진 영상 프레임과 활동 모델사이의 유사도를 최대화하는 하나의 활동 모델을 인식된 활동으로써 결정하는 단계;를 포함하는 것을 특징으로 하는 객체 활동 인식 방법.

## 【청구항 10】

제9항에 있어서, 상기 (d) 단계는,

$T$ 는 비디오 시퀀스를 구성하는 프레임의 수를 나타내는 양의 정수,  $Z_1, Z_1, \dots, Z_T$ 는 각각  $1, 2, \dots, T$  번째 프레임의 특징 벡터들이라 할 때, 영상 프레임  $O = \{Z_1, Z_1, \dots, Z_T\}$ 이 주어지면,  $E$ 는 상태 모델의 수라 할 때, 주어진 활동 모델 사전  $\{\lambda_1, \lambda_2, \dots, \lambda_E\}$ 으로부터 유사도  $P(O|\lambda)$ 를 최대화하는 하나의 활동 모델을 찾는 단계;를 포함하는 것을 특징으로 하는 객체 활동 인식 방법.

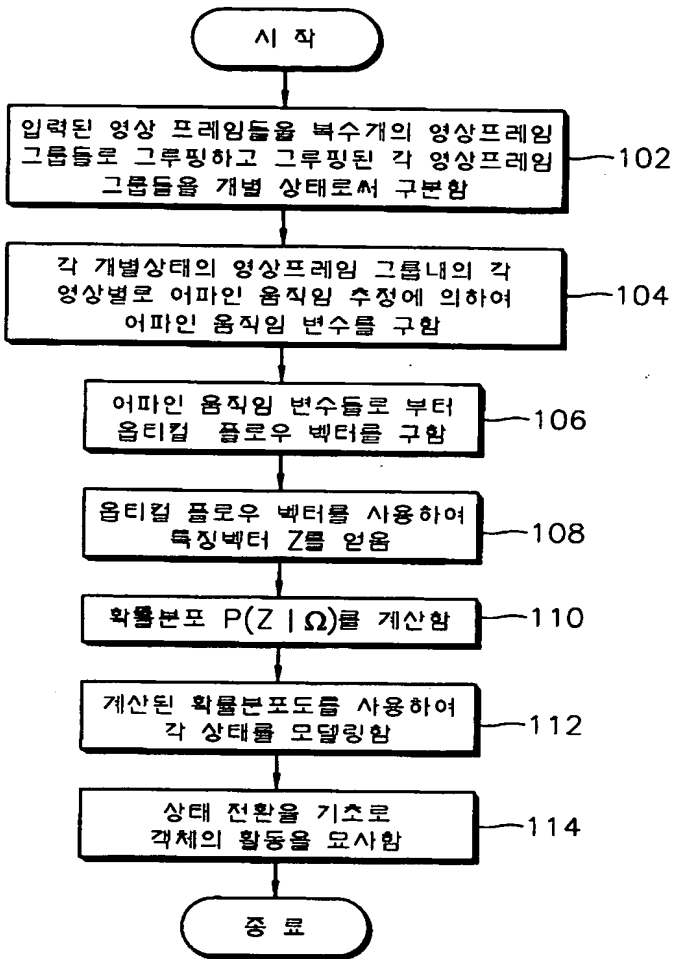
## 【청구항 11】

제10항에 있어서, 상기 전환 행렬은,

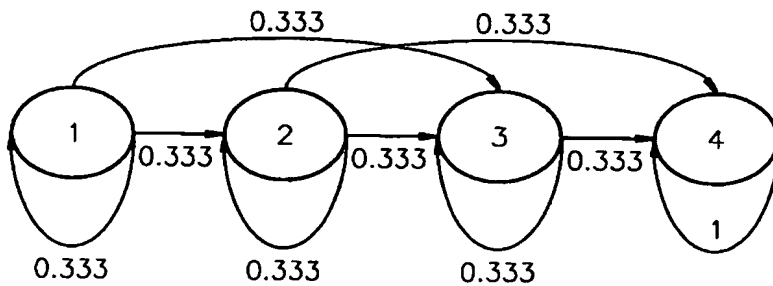
트레이닝 과정에서 장면  $j$ 에 해당하는 관측 심볼 확률  $B = \{b_j(\cdot)\}$ 을 기초로 EM(expectation-maximization) 알고리즘을 사용하여 얻어지는 것을 특징으로 하는 객체 활동 인식 방법.

## 【도면】

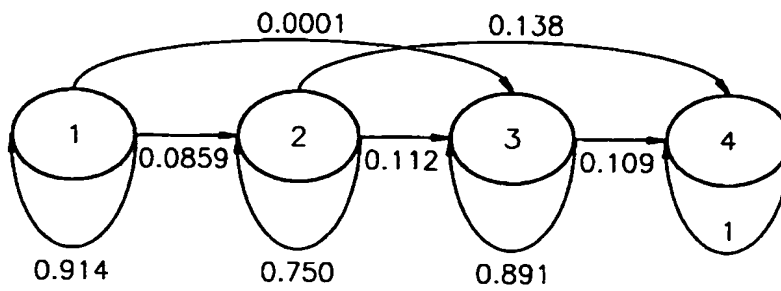
【도 1】



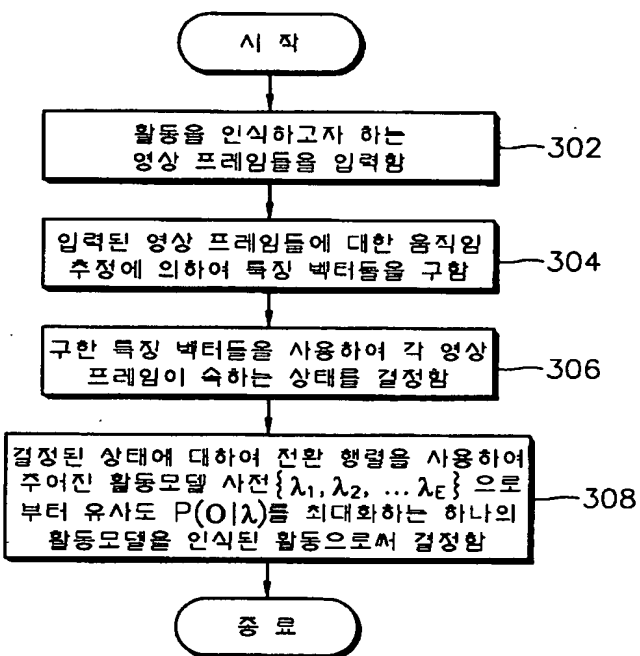
【도 2a】



【도 2b】



【도 3】



【서류명】	서지사항 보정서
【수신처】	특허청장
【제출일자】	2001.02.14
【제출인】	
【명칭】	삼성전자 주식회사
【출원인코드】	1-1998-104271-3
【사건과의 관계】	출원인
【제출인】	
【명칭】	더 리전트 오브 더 유니버시티 오브 캘리포니아
【출원인코드】	5-1999-020685-1
【사건과의 관계】	출원인
【대리인】	
【성명】	이영필
【대리인코드】	9-1998-000334-6
【포괄위임등록번호】	1999-009556-9
【포괄위임등록번호】	2000-055014-6
【사건의 표시】	
【출원번호】	10-2000-0086284
【출원일자】	2000.12.29
【심사청구일자】	2000.12.29
【발명의 명칭】	객체 활동 모델링 방법
【제출원인】	
【접수번호】	1-1-00-0287217-90
【접수일자】	2000.12.29
【보정할 서류】	특허출원서
【보정할 사항】	
【보정대상 항목】	발명자
【보정방법】	정정
【보정내용】	
【발명자】	
【성명의 국문표기】	최양림
【성명의 영문표기】	CHOI, Yang Lim
【주민등록번호】	710120-1830615



【우편번호】	442-190
【주소】	경기도 수원시 팔달구 우만동 105 우만 선경아파트 102동 11 12호
【국적】	KR
【발명자】	
【성명의 국문표기】	유운주
【성명의 영문표기】	YU, Yun Ju
【주민등록번호】	710725-1836311
【우편번호】	609-391
【주소】	부산광역시 금정구 장전1동 130-23 18/3
【국적】	KR
【발명자】	
【성명의 국문표기】	만주나스 방가로 에스.
【성명의 영문표기】	MANJUNATH, Bangalore S.
【주소】	미합중국 캘리포니아 93106-9560 산타바바라, 유니버시티 오브 캘리포티아
【국적】	IN
【발명자】	
【성명의 국문표기】	첸 칭-웨이
【성명의 영문표기】	CHEN, Ching-Wei
【주소】	미국 캘리포니아 93117 아이슬라비스타 6503 마드리드 알디 아파트 제4호
【국적】	US
【발명자】	
【성명의 국문표기】	선 신딩
【성명의 영문표기】	SUN, Xinding
【주소】	미합중국 캘리포니아 93106-9560 산타바바라, 유니버시티 오브 캘리포티아
【국적】	CN
【취지】	특허법시행규칙 제13조의 규정에 의하여 위와 같이 제출합니다. 대리인 이영필 (인)

1020000086284

출력 일자: 2001/2/22

【수수료】

【보정료】 0 원

【기타 수수료】 원

【합계】 0 원